

Análisis del Aprendizaje usando el Nivel de la Incertidumbre en la Adquisición de Conocimiento

Constanza R. Huapaya, Francisco A. Lizarralde, Graciela M. Arona, Jorge R. Vivas

Grupo de Investigación en Inteligencia Artificial (integrante del CIMEPB)

Departamento de Matemática-Facultad Ingeniería

Universidad Nacional de Mar del Plata

Juan B. Justo 4302. 7600. Mar del Plata

223-4816600 int 259

constanza.huapaya@gmail.com

francisco.lizarralde@gmail.com

grarona@fi.mdp.edu.ar

jvivas@mdp.edu.ar

Resumen

Como continuación de la investigación sobre el uso de la Minería de Datos Educacional (MDE) para descubrir nuevo conocimiento basado en datos de la actividad del estudiante se ha elegido estudiar la **incertidumbre** en la medición de la adquisición de conocimiento a fin de mejorar la estimación de perfiles cognitivos del estudiante sobre los cuales se está investigando desde hace dos años.

El conocimiento adquirido por los estudiantes es, en general, imperfecto. El grado de profundidad en la adquisición de habilidades cognitivas varía durante el aprendizaje. Se ha desarrollado un modelo difuso para medir la incertidumbre en la adquisición del conocimiento de los estudiantes.

En base a los seis objetivos educacionales de la taxonomía de Bloom revisada (que han sido llevados a tres niveles) se han definido conjuntos difusos asociados a cada uno de ellos. Se estudia la posibilidad de los perfiles de cada

estudiante en cada uno de los tres niveles propuestos. En base a esa información se puede calcular la incertidumbre posibilística total a fin de analizar el nivel adquisición de conocimiento de grupos de estudiantes.

La plataforma que usamos es un Ambiente Virtual de Aprendizaje. Con esos datos podemos monitorear características de los cursos de la plataforma.

Palabras clave: incertidumbre, modelos difusos, minería de datos educativos.

Contexto

La investigación descripta en esta presentación se desarrolla en el marco del Proyecto “Ambientes virtuales de aprendizaje para la enseñanza de la ingeniería” perteneciente al Departamento de Matemática de la Facultad de Ingeniería (integrante del Centro de Investigación en Procesos Básicos, Metodologías y Educación (CIMEPB, Facultad

de Psicología UNMDP) y aprobado por la Universidad Nacional de Mar del Plata.

Introducción

Generalmente las representaciones del mundo real son imperfectas y esa imperfección se traslada a los sistemas de información. Podemos comprobar que cuando se trabaja con modelos del mundo real, generalmente lo hacemos con aproximaciones. En contraposición, puede considerarse que la información es perfecta cuando es consistente, precisa y cierta. El término **imperfección de los datos** se usa aquí como el concepto más general, que incluye a los demás (inconsistencia, imprecisión e incertidumbre)[1] [2].

Información imperfecta	Imprecisión	Varios mundos satisfacen la sentencia	
	Inconsistencia	Conclusión incoherente	
	Incertidumbre	vaguedad (=borrosidad)	
		ambigüedad	no-especificidad conflicto

Figura 1: taxonomía de la información imperfecta

En la figura 1 se aprecia una clasificación de la información imperfecta. La **imprecisión** y la **inconsistencia** son propiedades asociadas al contenido de una sentencia (sin importar si algún mundo real es compatible con la información) mientras que la **incertidumbre** nace de la falta de información del mundo real para decidir si la sentencia que constituye la información es verdadera o falsa. La incertidumbre es una propiedad de la relación entre la información y nuestro conocimiento de la realidad. En una visión restringida puede verse a la incertidumbre como una deficiencia de la información, mientras que la información se percibe como la capacidad de reducir la incertidumbre.

Las sentencias pueden verdaderas o falsas, pero la **incertidumbre** aparece cuando el conocimiento que se posee acerca del mundo real no nos permite decidir si la sentencia es

verdadera o falsa. Si la certidumbre es el conocimiento completo sobre el valor de verdad de los datos, la incertidumbre es el conocimiento parcial sobre el valor de verdad de la información

La incertidumbre se refiere, entonces, a un estado de creencia parcial, modelado mediante el agregado a cada proposición (o evento o conjunto de posibles estados) de un peso que caracteriza nuestra confianza en la verdad de tal proposición (u ocurrencia del evento, o la afirmación de que el conjunto de estados posibles contiene al estado real de casos). La incertidumbre, además de originarse en la imprecisión, puede deberse a la inconsistencia o a la aleatoriedad.

Según Klir [3][4][5] existen dos grandes categorías de incertidumbre: la **vaguedad** y la **ambigüedad**. En general, la **vaguedad** emerge cuando es difícil encontrar límites precisos entre los mundos posibles, motivo por el cual se dificulta la elección del usuario sobre cuál de ellos elegir (*borrosidad*, *nebulosidad*, *difuso* aparecen como conceptos asociados a vaguedad). La **ambigüedad** se manifiesta cuando el usuario tiene dos o más posibilidades para elegir una opción, esto es, no puede distinguir una opción de otra. La ambigüedad está conectada a conceptos como *no-especificidad* y *relación uno-a-muchos*.

La **medida** de la vaguedad o borrosidad ha sido uno de temas asociados con el desarrollo de la teoría de **conjuntos difusos**. En este contexto existen en la literatura diversas medidas de borrosidad. En general, una medida de la borrosidad es una función

$$f : \mathcal{P}(X) \rightarrow \mathbb{R}$$

Donde $\mathcal{P}(X)$ es el conjunto de todos los subconjuntos de X y f asigna $f(A)$ a cada subconjunto difuso A de X caracterizando el grado de borrosidad de A .

La **medida** de la ambigüedad [6] vista como no-especificidad, disonancia y confusión ha sido estudiada en el contexto de las medidas de la plausibilidad y las medidas de creencia. En nuestro estudio nos interesa tratar con el principio de invariancia de la incertidumbre. En particular trabajamos con la incertidumbre posibilística total vista como suma de la no-especificidad y el conflicto.

Volviendo a nuestra investigación, se ha considerado muy pertinente el análisis de la **incertidumbre** asociada al proceso de enseñanza/ aprendizaje. En particular investigamos la relación entre la incertidumbre y el diagnóstico cognitivo [7,8, 9,10, 11].

En primer lugar, consideremos la evaluación subjetiva que llevan a cabo los docentes/evaluadores: se puede comprobar que existe el fenómeno de la **vaguedad** cuando se trata de analizar el nivel de adquisición del conocimiento por parte de los alumnos (es difícil precisar ese nivel de adquisición). Ha sido desarrollado [12,13] un modelo de diagnóstico basado en Lógica Difusa [14,15, 16] cuyo esquema se aprecia en la figura 2.

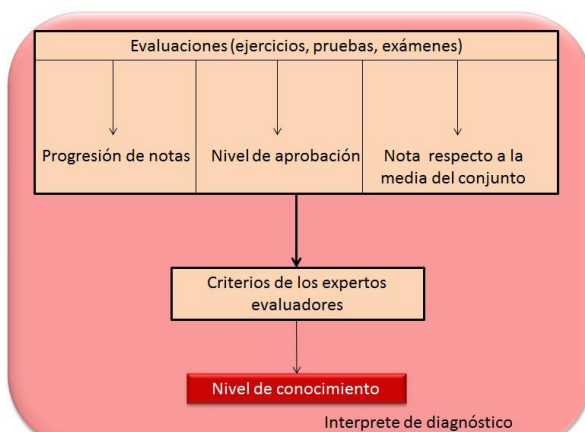


Figura 2: Modelo base del diagnóstico

En segundo lugar, examinemos características cognitivas comunes de los alumnos. A tal fin se han definido perfiles de aprendizaje de grupos de estudiantes usando el intérprete de diagnóstico de la figura 2. En este caso se

presenta el fenómeno de la **ambigüedad** (no es específico a que perfil pertenece cada alumno). Los perfiles están basados en la taxonomía revisada de Bloom [17,18], ver figura 3.

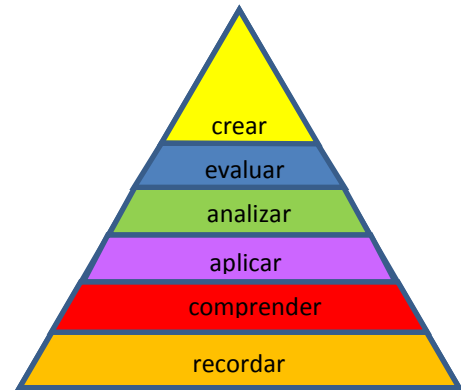


Figura 3: Taxonomía revisada de Bloom

A partir de la taxonomía se han construido tres niveles de complejidad creciente: A (recordar y comprender), B (aplicar y analizar) y C (evaluar y crear). Cada uno de estos niveles son valorados usando 5 variables lingüísticas para cada uno de los niveles (muy bajo (MB), bajo (B), medio (M), alto (Al) y muy alto (MA)). Ejemplo de un perfil es la terna (MA,M,B) indicando que el alumno ha alcanzado el nivel muy alto en el nivel A, medio en el nivel B y bajo en el nivel C. En la figura 4 se aprecia un esquema del cálculo de la incertidumbre en el aprendizaje de un grupo de estudiantes.

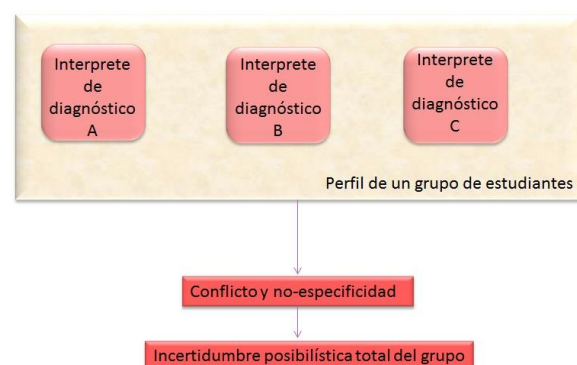


Figura 4: Modelo de medición de incertidumbre

Para construir, automáticamente, los perfiles se usa el Intérprete del Diagnóstico aplicado a cada uno de los niveles para cada uno de los

alumnos del grupo. Con esta información se estudia la posibilidad de todos los posibles perfiles para finalmente calcular la **incertidumbre posibilística total**. Cuanto más bajo sea este valor, más alto es el nivel de adquisición de conocimiento del grupo de estudiantes.

La presente investigación está enmarcada dentro de los métodos que estamos desarrollando para nuestra biblioteca de algoritmos propios de la Minería de Datos Educacional [19,20,21,22]. Ya están desarrolladas técnicas de clustering difuso y modelos difusos para el análisis de adquisición de conocimiento.

Líneas de investigación y desarrollo

- *Uso de un modelo difuso para analizar la incertidumbre en el proceso de aprendizaje de un grupo de estudiante.*
- *Análisis de la incertidumbre en la construcción de perfiles de los estudiantes.*
- *Construcción de un perfil cognitivo del estudiante basado en los datos filtrados por técnicas de Minería de Datos Educacional.*

Resultados y Objetivos

- Se ha desarrollado un modelo de perfiles cognitivos considerando una variable lingüística de salida (nivel de conocimiento) y tres variables lingüísticas de entrada (progresión de notas, nivel de aprobación de las pruebas y nota final respecto a la media del curso). El modelo se completa con 27 reglas difusas que capturan la experticia de los profesores. Este modelo ha sido encapsulado como un intérprete de diagnóstico.

- Se han desarrollado representaciones matemáticas del proceso de aprendizaje basadas en la taxonomía de Bloom a fin de afianzar el perfil cognitivo individual.
- Se ha desarrollado un espacio tridimensional que se genera a partir de la evaluación de tres habilidades cognitivas de cada estudiante: comprensión del problema, planteo general de la solución y eficiencia de la solución. Posteriormente, se aplicaron diversas técnicas de clustering y se analizaron los resultados a fin de encontrar los perfiles del grupo.
- Se ha desarrollado una técnica de medida de la incertidumbre para un grupo de estudiantes.

Formación de Recursos Humanos

El equipo de trabajo está integrado por especialistas en Informática, en Psicología y estudiantes de Ingeniería y Psicología. Actualmente se están desarrollando tres tesis de la Maestría en Tecnología Informática Aplicada a Educación de la Facultad de Informática de la Universidad Nacional de La Plata y una tesis de doctorado de la misma Facultad.

Referencias

- [1] Smets Ph.(1997) Imperfect information : Imprecision - Uncertainty. Uncertainty Management in Information Systems. From Needs to Solutions. A. Motro and Ph. Smets (eds.), Kluwer Academic Publishers. 225-254.
- [2] Smets Ph.(1998) Probability, Possibility, Belief: Which and Where ? Handbook of Defeasible Reasoning and Uncertainty Management Systems. Gabbay D. and Smets Ph. (Series Eds). Ph. Smets (Vol. eds.), Vol. 1: Quantified Representation of Uncertainty & Imprecision, Kluwer, Dordrecht 1-24.
- [3] Klir G. J. (1995) Principles of uncertainty: What are they? Why do we need them? Fuzzy Sets and Systems. (74)15-31.
- [4] Klir G.J. (1987) Where do we stand on measures of uncertainty, ambiguity, fuzziness and the like?. Fuzzy Stes and Systems . (24),141-160.

- [5] Klir G.J. (2006) Uncertainty and Information. Foundations of Generalized Information Theory . John Wiley & Sons.
- [6] Shafer G. (1976). A Mathematical Theory of Evidence. Princeton Univ. Press. Princeton.
- [7] Ma.J. y Zhou D. (2000) Fuzzy Sets Approach to the Assessment of Student-centered Learning. IEEE Transactions on Education, 43(2), 237-241.
- [8] Jeong H. y Biswas G. (2008). Mining Student Behavior Models in Learning-by-Teaching Environments. The 1st International Conference on Educational Data Mining. 127-136.
- [9] Grigoriadou M., Kornilakis H., Papanikolaou K. y Magoulas G. (2002) Fuzzy Inference for Student Diagnosis in Adaptive Educational Hypermedia. Methods and Applications of Artificial Intelligence. Lecture Notes in Computer Science, , Volume 2308/2002.
- [10] Stathacopoulou R., Grigoriadou M., Samarakou M., Mitropoulos D.(2007) Monitoring students' actions and using teachers' expertise in implementing and evaluating the neural network-based fuzzy diagnostic model. Expert Systems with Applications 32 , 955–975.
- [11]Chenn-Jung Huang, Ming-Chou Liu, San-Shine Chu, Chih-Lun Cheng.(2007). An intelligent learning diagnosis system for Web-based thematic learning platform. Computers & Education 48 , 658–679
- [12] Huapaya C.R.,Lizarralde F.A.Arona G.M. (2012) Modelo basado en lógica difusa para el diagnóstico cognitivo del estudiante. Formación universitaria. 5(1), 13-20.
- [13] Huapaya C.R. (2012) Proposal of Fuzzy Logic-based Students' learning assessment model. CACIC 2012 WTIAE. Bahía Blanca.
- [14] Zadeh L.A (1965). Fuzzy sets. Information and control 8. p 338-353
- [15] Zadeh L.A.. The concept of a linguistic variable and its application to approximate reasoning . I-III Information Sci 8, , 199-250 y 301-357. 1975
- [16] Dubois D. y Prade H. What are fuzzy rules and how to use them. Fuzzy sets and systems, 84:169-185. 1996.
- [17] Pohl, M. (2000). Bloom's (1956) Revised Taxonomy. Retrieved April 14, 2008, from the QSITE Higher Order Thinking Skills Online Course Web site: http://www.odu.edu/educ/roverbau/bloom/blooms_taxonomy.htm. Consulta: 8/3/2013.
- [18] Anderson, Et Al. (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives (Complete ed.). New York: Longman.
- [19] Hammouda, K., Kamel, M. (2006). Data Mining in e-learning. E-Learning Networked Environments and Architectures: A Knowledge Processing Perspective, Samuel Pierre (Ed.), Springer Book Series: Advanced Information and Knowledge Processing (pp. 1-28).
- [20] Romero C, Ventura S. y García E. (2008) *Data Mining in course management systems: Moodle case study and tutorial*. Computers & Education. Volume: 51, (1), 368-384
- [21] Romero, C., Ventura, S., & Bra, P. D. (2004). Knowledge discovery with genetic programming for providing feedback to courseware author. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, 14(5), 425–464.
- [22]Romero, C., & Ventura, S. (2006). Data mining in e-learning, Southampton, UK: Wit Press.